

Visually optimal DCT quantization matrices for individual images.

Andrew B. Watson

NASA Ames Research Center

beau@vision.arc.nasa.gov

1 Introduction

The JPEG, MPEG, and CCITT H.261 image compression standards employ the Discrete Cosine Transform as a basic mechanism (Wallace, 1991). In JPEG, the DCT is applied to 8 by 8 pixel blocks, followed by uniform quantization of the DCT coefficient matrix. The quantization bin-widths for the various coefficients are specified by a quantization matrix (QM). The QM is not defined by the standard, but is supplied by the user and stored or transmitted with the compressed image.

The principle that should guide the design of a QM is that it provide optimum visual quality for a given bit rate. QM design thus depends upon the visibility of quantization errors at the various DCT frequencies. In recent papers, Peterson *et al.* (Peterson, 1992; Peterson, Peng, Morgan, & Pennebaker, 1991) have provided measurements of threshold amplitudes for DCT basis functions at one viewing distance and several mean luminances, and Ahumada and Peterson (Ahumada, & Peterson, 1992) have devised a model that generalizes these measurements to other luminances and viewing distances. From this model, a

matrix can be computed which will insure that all quantization errors are below threshold. I call this the *image-independent perceptual* (IIP) approach.

While a great advance over the *ad hoc* matrices that preceded it, the IIP approach has several shortcomings. The fundamental limitation is that the matrix is computed independently of the image. This would not be a problem if visual thresholds for artifacts were fixed and independent of the image upon which they are superimposed, but unfortunately this is not the case.

First, visual thresholds increase with background luminance. The formula of Ahumada & Peterson describes the threshold for DCT basis functions as a function of a mean luminance. This would normally be taken as the mean luminance of the display. But variations in local mean luminance within the image will in fact produce substantial variations in DCT thresholds from block to block. I call this *luminance masking*.

Second, threshold for a visual pattern is typically reduced in the presence of other patterns, particularly those of similar spatial frequency and orientation, a phenomenon usually called *contrast masking*. This means that threshold error in a particular DCT coefficient in a particular block of the image will be a function of the value of that coefficient in the original image.

Third, the IIP approach ensures that any single error is below threshold. But in a typical image there are many errors, of varying magnitudes. The visibility of this error ensemble is not generally equal to the visibility of the largest error, but reflects a pooling of errors, over both frequencies and blocks of the image. I call this *error pooling*.

Fourth, when all errors are kept below a perceptual threshold a certain bit rate will result. The IIP method gives no guidance on what to do when a lower bit rate is desired. The ad hoc "quality factors" employed in some JPEG implementations, which usually do no more than multiply the quantization matrix by a scalar, will allow an arbitrary bit rate, but do not guarantee (or even suggest) optimum quality at that bit rate. I call this the problem of *selectable quality*.

Here I present a general method of designing a custom quantization matrix tailored to a particular image. This *image-dependent perceptual* (IDP) method incorporates solutions to each of the problems described above: luminance masking, contrast masking, error pooling, and selectable quality.

The solution is to develop a very simple model of perceptual error, based upon DCT coefficients, and to iteratively estimate the quantization matrix which yields a designated perceptual error for each coefficient.

2 Image-Independent Perceptual DCT Compression

Before considering the image-dependent method, we briefly review the image-independent approach. In the JPEG standard, the image is first divided into blocks of size {8,8}. Each block is transformed into its DCT, which we write c_{ijk} , where the indices $\{i,j\}$ index the DCT frequency (or basis function), and k indexes a block of the image. Though the blocks themselves form a two dimensional array, for present purposes a one dimensional index is sufficient. Each block is then quantized by dividing it, coefficient by coefficient, by a quantization matrix q_{ij} , and rounding to the nearest integer

$$u_{ijk} = \text{Round}\left[c_{ijk}/q_{ij}\right].$$

The quantization error e_{ijk} in the DCT domain is then

$$e_{ijk} = c_{ijk} - u_{ijk} .$$

For each frequency $\{i,j\}$ it is possible to measure psychophysically the smallest coefficient that yields a visible signal. We call this threshold t_{ij} . The maximum possible quantization error e_{ijk} is $q_{ij}/2$. The image-independent approach of Ahumada & Peterson (1992) ensures that all errors are invisible (below threshold), by setting

$$q_{ij} = 2 t_{ij} .$$

3 Luminance Masking

Detection threshold for a luminance pattern typically depends upon the mean luminance of the local image region: the brighter the background, the higher the luminance threshold (van Ness, & Bouman, 1967). This is usually called "light adaptation," but here we call it "luminance masking" to emphasize the similarity to contrast masking, discussed in the next section.

The effect of mean luminance upon the DCT thresholds is complex, involving both vertical and horizontal shifts of the contrast sensitivity function. We can compute a luminance-masked threshold matrix for each block in either of two ways. The first is to make use of a formula such as that supplied by Ahumada and Peterson (Ahumada, et al., 1992) , which we write symbolically as

$$t_{ijk} = ap[c_{00k}, i, j]$$

where c_{00k} is the DC coefficient of the DCT for block k . This solution is as complete and accurate as the underlying formula, but may be rather expensive to

compute. For example, in the *Mathematica* language, using a compiled function, and running on a SUN Sparc 2, it takes about 1 second per block.

A second, simpler solution is to approximate the dependence of t_{ij} upon c_{00k} with a power function:

$$t_{ijk} = t_{ij} (c_{00k} / \bar{c}_{00})^{a_r}$$

where \bar{c}_{00} is an average of the DC terms for the complete image, or more simply a nominal average of 1024 for an 8 bit image. The initial calculation of t_{ij} should be made assuming a luminance corresponding to \bar{c}_{00} (typically half the maximum luminance). The parameter a_r takes its name from the corresponding parameter in the formula of Ahumada and Peterson, wherein they suggest a value of 0.649. Note that luminance masking may be suppressed by setting $a_r=0$. More generally, a_r controls the degree to which this masking occurs. Note also that the power function makes it easy to incorporate a non-unity display Gamma, by multiplying a_r by the Gamma exponent.

This power function approximation is quite accurate over an upper range of luminances (above about 10 cd m⁻²). Except for very dark sections of an image, this range should be adequate. The discrepancy is also greatest at the lowest frequencies, especially the DC term. This could be corrected by adopting a matrix of exponents, one for each frequency. But note that the discrepancy is a conservative one, that is the threshold changes less with block luminance than the model calls for.

4 Contrast Masking

Contrast masking refers to the reduction in the visibility of one image component by the presence of another. This masking is strongest when both

components are of the same spatial frequency, orientation, and location. Here we consider only masking within a block and a particular DCT coefficient (It is possible to extend these ideas to masking between DCT coefficients, and across DCT blocks). We employ a model of visual masking that has been widely used in vision models, based on seminal work by Legge and Foley (Legge, & Foley, 1980). Given a DCT coefficient c_{ijk} and a corresponding absolute threshold t_{ijk} our masking rule states that the masked threshold m_{ijk} will be

$$m_{ijk} = \text{Max} \left[t_{ijk}, \left| c_{ijk} \right|^{w_{ij}} t_{ijk}^{1-w_{ij}} \right]$$

where w_{ij} is an exponent that lies between 0 and 1. Note that when $w_{ij}=0$, no masking occurs, and the threshold is constant at t_{ijk} . When $w_{ij} = 1$, we have what is usually called "Weber Law" behavior, and threshold is constant (for $c_{ijk} > t_{ijk}$) in log or percentage terms. A typical empirical value is 0.7. Because the exponent w_{ij} may differ for different frequencies, we allow a matrix of exponents equal in size to the QM ({8,8}).

Because the effect of the DC coefficient upon thresholds has already been expressed by luminance masking, we specifically exclude the DC term from the contrast masking, by setting the value of $w_{00} = 0$. It is interesting that while contrast masking is assumed to be independent from coefficient to coefficient (frequency to frequency), in the case of luminance masking the DC frequency affects *all* other frequencies.

5 Perceptual Error and Just-Noticeable-Differences

In vision science, we often express the magnitude of a signal in multiples of the threshold for that signal. These threshold units are often called "just-

noticeable differences," or *jnd*'s. Having computed a masked threshold m_{ijk} , the error DCT may therefore be expressed in *jnd*'s as

$$d_{ijk} = e_{ijk} / m_{ijk}$$

Each value of d_{ijk} is an error in a particular frequency and block, expressed as a proportion of the just-detectable error in that frequency and block. Thus all the errors are now in the "common coin" of perceptual error, the *jnd*.

6 Error Pooling

To pool the errors in the *jnd* DCT we employ another standard feature of current vision models: the so-called β -norm (or Minkowski metric). It often arises from an attempt to combine the separate probabilities that individual errors will be seen, in the scheme known as "probability summation" (Graham, 1977 ; Robson, & Graham, 1981; Watson, 1979) . We pool the *jnds* for a particular frequency $\{i,j\}$ over all blocks k as

$$p_{ij} = \left(\sum_k |d_{ijk}|^\beta \right)^{1/\beta}$$

Different values of the exponent β implement different types or degrees of pooling. When $\beta=1$, the pooling is linear summation of absolute values. When $\beta=2$, the errors combine quadratically, in what may be thought of as a measure of energy or standard deviation. When $\beta=\infty$ (in practice, a large number such as 100 will do), the pooling rule becomes a maximum-of operation: only the largest error matters. In psychophysical experiments that examine summation over space, a β of between 3 and 4 has been observed (Robson, et al., 1981). The exponent β is given here as a scalar, but may of course be made a matrix equal in size to the QM to allow differing pooling behavior for different DCT frequencies.

This matrix p_{ij} of "pooled jnds" is now a simple measure of the visibility of artifacts within each of the frequency bands defined by the DCT basis functions. I call it the "perceptual error matrix."

7 Optimization Rule

This perceptual error matrix p_{ij} could be used for many purposes, but here we use it to compute a quantization matrix by means of a simple rule. We attempt to set each entry in p_{ij} to a constant target value ψ . The intuition here is that each entry in the matrix p_{ij} corresponds (at least monotonically) with the visibility of a particular class of artifact: that of the corresponding frequency (basis function). We assume that these visibilities are independent. We therefore attempt to set all artifact classes to a particular level of visibility. To do otherwise would be to favor or penalize one sort of error over another, which would be inappropriate since all are expressed in common units of visibility. It can be shown that this rule is equivalent to assuming that the pooling exponent *within* the perceptual error matrix is infinite.

8 Results

Using the rule described above and various standard optimization methods, it is possible to calculate for a given image either the quantization matrix that yields a particular perceptual error, or which minimizes the perceptual error for a given bit-rate. We have applied these techniques to a number of standard digital images and find that they yield consistently better apparent quality than standard techniques. One example is given in Fig. 1.

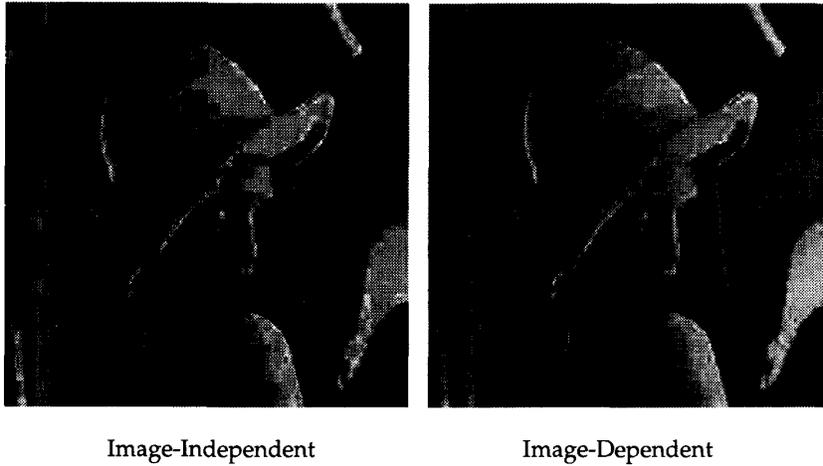


Figure 1. The Lena image (Weber, 1983) of 256^2 pixels compressed to 0.25 bit/pixel using image-independent and image-dependent JPEG techniques.

From the standpoint of computational complexity, this algorithm adds only a modest amount to the cost of JPEG image compression. All optimization takes place in the DCT domain, so no additional forward or inverse DCT's are required. The DCT mask is computed only once, and consists of a few calculations on each DCT pixel. The estimation of the quantization matrix requires a maximum of ten (and probably many fewer) iterations, each of which consists of a modest number of simple operations on each DCT pixel. It is certainly a smaller burden than requiring the user to repeatedly compress, reconstruct, and visually assess the result.

Acknowledgments

I thank Albert Ahumada, Jr. and Heidi Peterson for useful discussions. This work supported by NASA RTOP 506-71-51.

References

- Ahumada, A. J., Jr., & Peterson, H. A. (1992). Luminance-Model-Based DCT Quantization for Color Image Compression. In B. E. Rogowitz (Ed.), Human Vision, Visual Processing, and Digital Display III (pp. 365-374). Proc. of the SPIE.
- Graham, N. (1977). Visual detection of aperiodic spatial stimuli by probability summation among narrowband detectors. Vision Res. 17, 37-652 .
- Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. Journal of the Optical Society of America 70(12), 1458-1471.
- Peterson, H. A. (1992). DCT basis function visibility in RGB space. In J. Morreale (Ed.), Society for Information Display Digest of Technical Papers Playa del Rey, CA: Society for Information Display.
- Peterson, H. A., Peng, H., Morgan, J. H., & Pennebaker, W. B. (1991). Quantization of color image components in the DCT domain. SPIE, 210-222.
- Robson, J. G., & Graham, N. (1981). Probability summation and regional variation in contrast sensitivity across the visual field. Vision Res. 21, 409-418.
- van Ness, F. L., & Bouman, M. A. (1967). Spatial modulation transfer in the human eye. Journal of the Optical Society of America 57, 401-406.
- Wallace, G. (1991). The JPEG still picture compression standard. Communications of the ACM 34(4), 30-44.
- Watson, A. B. (1979). Probability summation over time. Vision Res. 19, 515-522.
- Weber, A. (1983). Image data base (USCIPI Report 1070) . Los Angeles, CA: Image Processing Institute, University of Southern California.
-